GCCAGGTCTGGCACCATGCACTAGGATACCCAGAACGCTGCAAGGCCACGCC
CTCCTCACTTCAGGGGTCACTCTCCCCATTGCCCACCACCCCACCATGGCTGGG
GATCGGCTCCCGAGGAAGGTGATGGACGCAAAGAAACTGGCCAGCCTGCTGC
GTGGCGGGCCTGGGGGACCCTTGGTCATCGACAGCCGGTCCTTCGTGGAGTAT
AACAGCTGCCACGTGCTGAGCTCTGTGAATATCTGCTGTTCAAAGCTGGTGAA
GCGGCGCCTTCAGCAGGGAAAAGTGACAATTGCTGAGCTTATCCAGCCTGCTA
CACGGAGCCAGGTGGATGCCACAGAACCACAGGATGTAGTGGTGTATGACCA
GAGCACACGAGATGCCAGCGTGCTGGCAGCAGACAGCTTCCTGTCCATCCTGC
TCAGCAAGCTGGACGGCTGCTTCGACAGTGTGGCCATCCTCACAGGAGGCTTC
GCCACCTTCTCCTCCTGCTTCCCTGGCCTCTGTGAGGGCAAGCCTGCCACTCTA
CCGTCCATGAGCCTCTCTCAGCCCTGCCTGCCTGTGCCCAGTGTTGGCCTGACC
CGAATCCTGCCTCACCTCTACCTGGGCTCTCAGAAAGATGTCTTGAACAAGGA
TCTGATGACCCAAAACGGAATAAGCTATGTCCTCAATGCCAGCAACTCCTGCC
CTAAACCGGACTTCATCTGTGAGAGCCGTTTCATGCGTATCCCCATCAATGAC
AACTACTGTGAAAAGCTGCTGCCCTGGCTGGACAAGTCCATCGAGTTTATTGA
TAAAGCCAAGCTGTCCAGCTGCCAAGTCATTGTTCACTGTCTGGCTGGCATCTC
TCGCTCTGCCACCATTGCCATCGCGTACATCATGAAAACCATGGGCATGTCTTC
TGACGACGCATACAGGTTTGTGAAGGATCGGCGCCCCTCCATCTCGCCCAACT
TCAACTTCCTGGGCCAGTTGCTGGAGTATGAGAGGAGTCTGAAGCTGCTGGCT
GCCCTGCAGACTGATGGACCTCACTTGGGGACCCCTGAGCCCCTCATGGGCCC
GGCAGCAGGCATCCCACTGCCCCGGCTGCCACCATCTACCTCAGAGAGCGCTG
CCACTGGGAGCGAGGCAGCCACCGCAGCCAGGGAGGGCAGCCCAAGTGCTGG
AGGGGATGCTCCGATCCCCAGCACAGCTCCAGCCACCAGCGCGCTGCAGCAG
GGCCTGCGTGGCCTGCACCTCTCCTCTGACCGCCTCCAGGACACCAACCGCCT
CAAGCGTTCCTTTTCCCTGGACATCAAGTCGGCCTATGCACCCAGCAGGAGGC
CCGACTTTCCCGGCCCACCCGACCCCGGTGAAGCCCCGAAGCTCTGCAAGCTG
GACAGCCCGTCTGGGGGCACACTGGGCCTGCCCTCGCCCAGCCCAGACAGCCC
GGACTCCGTTCCAGAGTGCCGCCCACGACCCCGCCGGCGACGCCCCCCGGCTA
GTTCGCCTGCCCGCTCCCCCGCGCATGGTCTGGGCCTGAACTTTGGAGACACG
GCCCGGCAGACTCCACGGCACGGCCTCTCGGCCCTGTCGGCGCCCGGGCTGCC
TGGCCCTGGCCAGCCGGCTGGCCCCGGGGGCTGGGTGCCGCCACTGGACTCCC
CAGGCACACCGTCGCCCGACGGCCCCTGGTGCTTCAGCCCCGAGGGCGCGCA
GGGTCCAGGCGCTGTGTTCTCCGCCTTTGGCCGGGTAAGTGCAGGCGCACCTG
GACCCGGTAACAGCAGCAGCAGCGGTGGTGGTGGTGGTGGTGGCGGCGG
CGGCGGCGGCGGCGGCGGCGGCAGCAGCAGCAGCAACAGCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGTAGTAGTAGTAGTAGTGACCTG
CGGAGGCGGGATGTGCGGACCGGCTGGCCCGAGGAGCCTGCTGCAGATGCAC
AGTTCAAGAGGCGCAGCTGCCAGATGGAGTTCGAAGAGGGCATGGTGGAGGG
GCGGGCACGTGGCGAGGAGCTGGCAGCCCTGGGCAAGCAAACCAGCTTCTCT
GGCAGCGTGGAGGTCATCGAAGTATCGTGACCCTTCAGAAGTCCCTGTGCCCT
TGCTCCAGCCAGGCCAGGTATAAATATATATTATATATAAAACACACAGAAAA
GGTAAATGGTTTTACTGCAATTTTTATCAAGAAGTAAATATTTCGATTTTTTAT
TTATTTAAGCTAGTGATCTGGCAACTGTGCGGGGCGGCCCTAAAGCTCTGTTTT
TACTGTCTGGTATTTAAACTGAAACAGGTTTCTAAGCAATATGAGGCCACCTT
CAATCCCAAACTGGGTTGACAGGCCTGGGCCCCTCCTTGCCCCTCCCCTCTGG
AAACATTACTGACCTTTCAAAGAGCTGCCCAGCTTTCCTGCACTTTTTACATAA
GAAAAAGGGGGGGGGGGAA   (SEQ ID NO:1)

FIGURE 1

underlined = deleted in targeting construct

[ ] = sequence flanking Neo insert in targeting construct

[GCCAGGTCTGGCACCATGCACTAGGATACCCAGAACGCTGCAAGGCCACGCCCTCCTCAC
TTCAGGGGTCACTCTCCCCATTGCCCACCACCCCACCATGGCTGGGGATCGGCTCCCGAG
GAAGGTGATGGACGCAAAGAAA]CTGGCCAGCCTGCTGCGTGGCGGGCCTGGGGGACCCTT
GGTCATCGACAGCCGGTCCTTCGTGGAGTATAACAGCTGCCACGTGCTGAGCTCTGTGAA
TATCTGCTGTTCAAAGCTGGTGAAGCGGCGCCTTCAGCAGGGAAAAGTGACAATTGCTGA
GCTT[ATCCAGCCTGCTACACGGAGCCAG]GTGGATGCCACAGAACCACAGGATGTAGTGGT
GTATGACCAGAGCACACGAGATGCCAGCGTGCTGGCAGCAGACAGCTTCCTGTCCATCCT
GCTCAGCAAGCTGGACGGCTGCTTCGACAGTGTGGCCATCCTCACAGGAGGCTTCGCCAC
CTTCTCCTCCTGCTTCCCTGGCCTCTGTGAGGGCAAGCCTGCCACTCTACCGTCCATGAG
CCTCTCTCAGCCCTGCCTGCCTGTGCCCAGTGTTGGCCTGACCCGAATCCTGCCTCACCT
CTACCTGGGCTCTCAGAAAGATGTCTTGAACAAGGATCTGATGACCCAAAACGGAATAAG
CTATGTCCTCAATGCCAGCAACTCCTGCCCTAAACCGGACTTCATCTGTGAGAGCCGTTT
CATGCGTATCCCCATCAATGACAACTACTGTGAAAAGCTGCTGCCCTGGCTGGACAAGTC
CATCGAGTTTATTGATAAAGCCAAGCTGTCCAGCTGCCAAGTCATTGTTCACTGTCTGGC
TGGCATCTCTCGCTCTGCCACCATTGCCATCGCGTACATCATGAAAACCATGGGCATGTC
TTCTGACGACGCATACAGGTTTGTGAAGGATCGGCGCCCCTCCATCTCGCCCAACTTCAA
CTTCCTGGGCCAGTTGCTGGAGTATGAGAGGAGTCTGAAGCTGCTGGCTGCCCTGCAGAC
TGATGGACCTCACTTGGGGACCCCTGAGCCCCTCATGGGCCCGGCAGCAGGCATCCCACT
GCCCCGGCTGCCACCATCTACCTCAGAGAGCGCTGCCACTGGGAGCGAGGCAGCCACCGC
AGCCAGGGAGGGCAGCCCAAGTGCTGGAGGGGATGCTCCGATCCCCAGCACAGCTCCAGC
CACCAGCGCGCTGCAGCAGGGCCTGCGTGGCCTGCACCTCTCCTCTGACCGCCTCCAGGA
CACCAACCGCCTCAAGCGTTCCTTTTCCCTGGACATCAAGTCGGCCTATGCACCCAGCAG
GAGGCCCGACTTTCCCGGCCCACCCGACCCCGGTGAAGCCCGAAGCTCTGCAAGCTGGA
CAGCCCGTCTGGGGGCACACTGGGCCTGCCCTCGCCCAGCCCAGACAGCCCGGACTCCGT
TCCAGAGTGCCGCCCACGACCCCGCCGGCGACGCCCCCCGGCTAGTTCGCCTGCCCGCTC
CCCCGCGCATGGTCTGGGCCTGAACTTTGGAGACACGGCCCGGCAGACTCCACGGCACGG
CCTCTCGGCCCTGTCGGCGCCCGGGCTGCCTGGCCCTGGCCAGCCGGCTGGCCCCGGGGG
CTGGGTGCCGCCACTGGACTCCCCAGGCACACCGTCGCCCGACGGCCCCTGGTGCTTCAG
CCCCGAGGGCGCGCAGGGTCCAGGCGCTGTGTTCTCCGCCTTTGGCCGGGTAAGTGCAGG
CGCACCTGGACCCGGTAACAGCAGCAGCAGCGGTGGTGGTGGTGGTGGTGGTGGCGGCGG
CGGCGGCGGCGGCGGCGGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG
CAGCAGCAGCAGCAGCAGTAGTAGTAGTAGTAGTGACCTGCGGAGGCGGGATGTGCG
GACCGGCTGGCCCGAGGAGCCTGCTGCAGATGCACAGTTCAAGAGGCGCAGCTGCCAGAT
GGAGTTCGAAGAGGGCATGGTGGAGGGGCGGGCACGTGGCGAGGAGCTGGCAGCCCTGGG
CAAGCAAACCAGCTTCTCTGGCAGCGTGGAGGTCATCGAAGTATCGTGACCCTTCAGAAG
TCCCTGTGCCCTTGCTCCAGCCAGGCCAGGTATAAATATATATTATATATAAAACACACA
GAAAAGGTAAATGGTTTTACTGCAATTTTTATCAAGAAGTAAATATTTCGATTTTTTATT
TATTTAAGCTAGTGATCTGGCAACTGTGCGGGGCGGCCCTAAAGCTCTGTTTTTACTGTC
TGGTATTTAAACTGAAACAGGTTTCTAAGCAATATGAGGCCACCTTCAATCCCAAACTGG
GTTGACAGGCCTGGGCCCCTCCTTGCCCCTCCCCTCTGGAAACATTACTGACCTTTCAAA
GAGCTGCCCAGCTTTCCTGCACTTTTTACATAAGAAAAAGGGGGGGGGGGAA
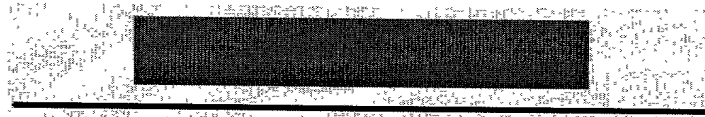
FIGURE 2A

Gene Sequence Structure *
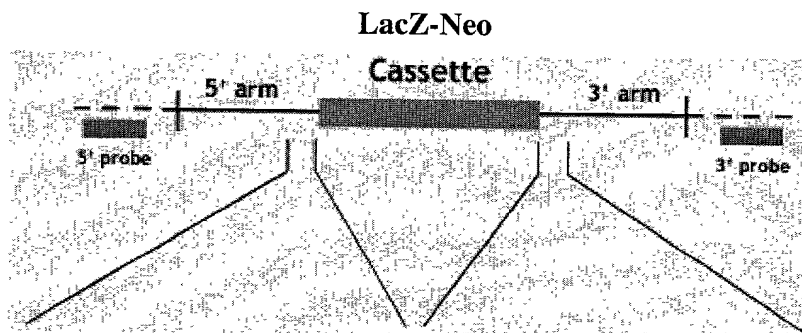
143 bp   Sequence Deleted   304 bp

Size of full-length cDNA: 2453 bp

Targeting Vector*
(genomic sequence)

LacZ-Neo

Arm Length:
5': 2.5 kb
3': 2 kb

Targeting Vector
Endogenous Locus

* Not drawn to scale

```
5'>TCCTGGGAGCCAGCTATAGCT
ACCCAGATCCCACCATCTGCTGAC
TATTCACCTTTCCCCCAGGTCTGG
CACCATGCACTAGGATACCCAGAA
CGCTGCAAGGCCACGCCCTCCTCA
CTTCAGGGGTCACTCTCCCCATTG
CCCACCACCCCACCATGGCTGGGG
ATCGGCTCCCGAGGAAGGTGATGG
ACGCAAAGAAA<3'
(SEQ ID NO:2)
```

```
5'>ATCCAGCCTGCTACACGAAGC
CAGGTACCTGTGGCCCACCCTTGC
ATGCGTCTTCAGGGCTGACCATTC
CTGAGCAAACAGACCTATGTCACC
TCTGAAAGAGACAGAGGAGCTCCC
AGGCCTGGTGCCAAGAGTCCTCTG
ATAAGGCATTTCCCCCTCGCTGTC
CCTCCGTTCCAAACAGGGTTCCTT
GGGGTCAGAGC<3'
(SEQ ID NO:3)
```

FIGURE 2B